

**CST1**  
**COMPUTER SCIENCE TRIPOS Part IB**

---

Thursday 6 June 2024      13:30 to 16:30

---

COMPUTER SCIENCE Paper 7

*Answer **five** questions.*

*Submit the answers in five **separate** bundles, each with its own cover sheet. On each cover sheet, write the numbers of **all** attempted questions, and circle the number of the question attached.*

**You may not start to read the questions  
printed on the subsequent pages of this  
question paper until instructed that you  
may do so by the Invigilator**

STATIONERY REQUIREMENTS

*Script paper*

*Blue cover sheets*

*Tags*

SPECIAL REQUIREMENTS

*Approved calculator permitted*

## 1 Artificial Intelligence

For a two-class classification problem with classes  $C_1$  and  $C_2$ , we will use the following linear classifier with probabilistic output

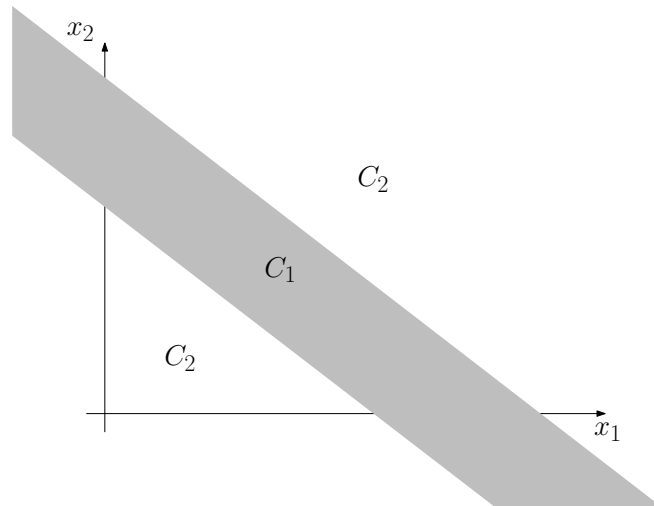
$$\Pr(C_1|\mathbf{x}, \mathbf{w}, w_0, \theta) = \sigma_\theta(\mathbf{w}^T \mathbf{x} + w_0) \quad (1)$$

where  $\mathbf{x}$  is an input vector and  $\mathbf{w}$ ,  $w_0$  and  $\theta$  are parameters, and

$$\sigma_\theta(x) = \frac{1}{1 + \exp(-\theta x)}$$

is the activation function.

- (a) The activation function  $\sigma_\theta$  often appears assuming  $\theta = 1$ . What role does  $\theta$  play when its value can be set freely? [1 mark]
- (b) You are provided with a collection  $\mathbf{s} = ((\mathbf{x}_1, y_1), \dots, (\mathbf{x}_m, y_m))$  of  $m$  training examples where  $y_i \in \{0, 1\}$  for  $i = 1, \dots, m$ . You suspect that the data has a band-like structure. In two dimensions:



Explain how the linear classifier can be modified, changing only its activation function and without adding further parameters, such that it can be trained on this type of data. It should be possible to set the width of the  $C_1$  region as part of the training. [6 marks]

- (c) When using a two-class classifier  $h(\mathbf{x}, \mathbf{p})$  that has parameters  $\mathbf{p}$  and outputs probabilities, one often uses the error measure

$$E(\mathbf{p}) = - \sum_{i=1}^m y_i \log h(\mathbf{x}_i, \mathbf{p}) + (1 - y_i) \log(1 - h(\mathbf{x}_i, \mathbf{p})).$$

Explain how the gradient descent method can be applied to solve this classification problem for the data described in Part (b), giving explicit formulae for the gradients. [13 marks]

## 2 Artificial Intelligence

This questions concerns the *Partial-Order Planning Algorithm*.

- (a) How does an *ordering constraint* differ from a *causal link*? [2 marks]
- (b) Why is it necessary for the Partial-Order Planning Algorithm to consider both ordering constraints and causal links? [2 marks]

A game involves filling a finite grid with coloured tiles. For a position  $(x, y)$  not on the first row or column, its *ancestors* are defined as positions  $(x - 1, y)$  and  $(x, y - 1)$ , where the origin is at the bottom-left. For a given grid state, a tile can be placed at an empty position  $(x, y)$  if and only if:

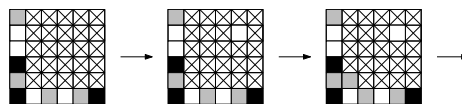
- its four adjacent squares are empty; or
- both ancestors have a tile, and the placement satisfies the Placement Rules; or
- only one ancestor has a tile, and the other ancestor can be filled later in a way consistent with the Placement Rules.

If a tile has been placed on every square, then the **Finish** state has been reached.

We will focus on a specific instance of the game with a 6-by-6 grid, where the colours are white, grey and black, and the Placement Rules are as follows:



In our instance the start state is fixed. Here are some possible moves from the start state, where a crossed square represents an empty grid position:



Solve the following problems using the Partial-Order Planning Algorithm.

- (c) Explain how the **Start** and **Finish** states can be represented for the instance given above. Only a description is required. [3 marks]
- (d) Explain how actions can be represented modeling the placement of tiles, by giving examples based on the specific instance described above. [7 marks]
- (e) Explain the concepts of *promotion* and *demotion*, which are of central importance in the Partial-Order Planning algorithm. To what extent might they be useful in the case of this particular algorithm and planning problem, given your formulation? Illustrate your answer with an example. [6 marks]

### 3 Economics, Law and Ethics

The new Master of (fictional) Porterhouse would like to modernise the college website, using generative AI to create both the html and content. They bring the proposal to the College Council, of which you are a member.

- (a) Discuss the challenges presented by intellectual property and generative AI. [6 marks]
- (b) Use either consequentialist or deontological reasoning to argue in favour of or against the proposal. Additionally, please define the ethical philosophy you will be employing in your argumentation. [6 marks]
- (c) One of the Fellows is concerned about the existential risk of generative AI. Discuss this risk and explain other risks from AI. [8 marks]

### 4 Economics, Law and Ethics

- (a) Explain information asymmetry, and how it can lead to a lemon market. [6 marks]
- (b) What solutions can help prevent a lemon market? [6 marks]
- (c) How is information asymmetry relevant to privacy? [8 marks]

## 5 Formal Models of Language

An intrepid linguist encounters an alien while exploring a remote planet, she documents an alien language that has four sounds,  $\Sigma = \{a, e, f, m\}$ . The following are examples of the alien's speech and a translation:

$e\ amaf$	nice spaceship
$e\ faem$	nice spaceboots
$e\ e\ amaf$	very nice spaceship
$e\ e\ e\ faem$	extremely nice spaceboots

The linguist hypothesises that Expression 1 matches sound sequences in the alien language.

Expression 1:  $e^*(a|e|f|m)(a|e|f|m)^*$

- (a) Describe the strings that are matched by Expression 1. [2 marks]
- (b) Provide a Finite State Automaton that can generate the language defined by Expression 1. [2 marks]

As the linguist continues to explore, she encounters a grumpy alien and documents the following utterances:

$fama\ e\ amaf$	not nice spaceship (unpleasant spaceship)
$meaf\ e\ e\ faem$	very unpleasant spaceboots
$afaf\ e\ e\ e\ fafa$	extremely unpleasant gift

The linguist hypothesises that Expression 2 matches sound sequences in the language.

Expression 2:  $w^{-1}e^*w$  where  $w \in \Sigma^*$

- (c) Can the linguist draw a Finite State Automaton to generate the language defined by Expression 2? Provide a proof for your answer. [5 marks]
- (d) Provide a grammar that can generate the language defined by Expression 2. [5 marks]

On the far side of the planet, the linguist encounters a new dialect and documents the following utterances:

$amaf\ e\ amaf$	unpleasant spaceship
$faem\ e\ e\ faem$	very unpleasant spaceboots
$mefaem\ e\ e\ e\ mefaem$	extremely unpleasant earthling

- (e) Provide a general expression to match such sound sequences. [1 mark]
- (f) Can you use a Context Free Grammar to generate the language defined by your expression? Provide a proof for your answer. [5 marks]

## 6 Formal Models of Language

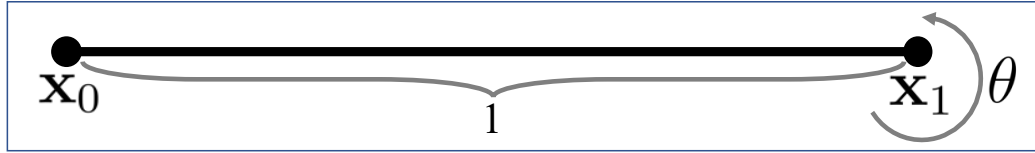
Two friends, Andrew and Brian, live on the top of adjacent hills. In order to reduce their telephone bills, they decide to try communicating using a flag. They use one flag that can be held in one of four positions (Left, Right, Up, Down). To evaluate their proposed system they do a trial. Every 5 seconds Andrew holds the flag in a different position and Brian notes down where he thinks the flag is. The results of their trial are below:

		ANDREW'S FLAG POSITION			
		LEFT	RIGHT	UP	DOWN
BRIAN'S OBSERVATION	LEFT	2	0	1	1
	RIGHT	2	0	1	1
	UP	2	8	4	2
	DOWN	2	0	2	4

- Let  $X$  be a random variable for the flag positions produced by Andrew. Assuming the trial is probabilistically representative of the underlying distribution, write down the distribution of  $X$  over the discrete set of flag events  $\mathcal{X} = \{\text{LEFT}, \text{RIGHT}, \text{UP}, \text{DOWN}\}$  [1 mark]
- What is the entropy,  $H(X)$ , of  $X$  in bits? Provide relevant equations **AND** describe in words what  $H(X)$  represents in this scenario. [2 marks]
- Let  $Y$  be a random variable for the flag positions seen by Brian. Calculate the joint entropy  $H(X, Y)$  in bits **AND** comment on whether this is the ideal joint entropy for this scenario. Provide relevant equations. [2 marks]
- What is the conditional entropy  $H(Y|X)$  in bits? Provide relevant equations **AND** describe in words what  $H(Y|X)$  represents in this scenario. [2 marks]
- Calculate the mutual information between  $X$  and  $Y$  in bits **AND** explain what this tells us about their communication system. Provide relevant equations. [3 marks]
- Provide an attested lower bound on the channel capacity for Andrew and Brian's communication system. Explain your answer providing any relevant equations. [3 marks]
- Do you think this will be an effective way for Andrew and Brian to communicate? Explain your answer. [2 marks]
- Andrew suggests using two positions for the flags instead of four. What do you think about this suggestion? What two positions would you choose and why? Explain how you could quantify the benefit (a numerical answer is not required). [5 marks]

## 7 Further Graphics

- (a) You are given a unit length rod as in the figure. There is no rotation at one end of the rod at point  $\mathbf{x}_0$  and a rotation around the plane normal  $\mathbf{n}$  is defined at the other end  $\mathbf{x}_1$ . This rotation is interpolated along the rod. The interpolation weight for  $\mathbf{x}$  on the rod is  $\|\mathbf{x} - \mathbf{x}_i\|$  for the rotation at  $\mathbf{x}$ ,  $i = 0, 1$ .



- (i) Write the rotation in quaternion form at  $\mathbf{x}$  given a rotation of  $\theta \leq \pi$  at  $\mathbf{x}_1$  assuming shortest path interpolation of rotations in  $\text{SO}(3)$ . [3 marks]
  - (ii) At  $\mathbf{x}_1$  at time  $t = 0$  there is no rotation and at  $t = 1$  the rotation angle is  $\pi/2$ . Write the quaternion at  $\mathbf{x}$  at any time in  $[0, 1]$  assuming the shortest path interpolation in  $\text{SO}(3)$  over time and over the rod. [3 marks]
  - (iii) Answer (ii), this time assuming linear blending of quaternions along the rod and the shortest path interpolation in  $\text{SO}(3)$  over time. [3 marks]
  - (iv) Write an expression for the norm of the quaternion in (iii). [3 marks]
- (b) Given a triangle in 3D with vertex locations  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ ,
- (i) determine a condition on the vertices for the triangle to define a valid plane, [1 mark]
  - (ii) define a parametric form  $\mathbf{p}(u, v)$  for the plane of the triangle assuming it defines a valid plane, [2 marks]
  - (iii) write an expression for the normal of the plane, [1 mark]
  - (iv) write the steps of an algorithm to find the closest point of a point  $\mathbf{x}$  in space on the triangle. [4 marks]

## 8 Further Graphics

This question is on rendering a scene via ray tracing and the rendering equation.

- (a) You are tasked with rendering a video of a scene via ray tracing. Only the camera is moving, the rest of the scene is static, the light sources emit light of a single wavelength, and surfaces do not emit light. You are allowed to shoot one ray per pixel and do pre-computation for a rotation and position of the camera. For each of the following cases, explain if this is possible, and if so, describe what you would pre-compute and store in single-channel textures, and how you would finally render the scene.
- (i) Local illumination and a single BRDF (bidirectional reflectance distribution function) for all surfaces. [3 marks]
  - (ii) Global illumination and diffuse reflection for all surfaces. [3 marks]
- (b) We are rendering a scene via path tracing and approximating integrals with importance-sampling.
- (i) Assume you can importance-sample with the exact incoming radiance. Write the resulting approximation of outgoing radiance at any point and direction with  $N$  samples. What is the length (number of segments) of the light path to compute this approximation? Why is this length finite? [5 marks]
  - (ii) There is only a single point light source in a scene, only light paths of length three are allowed, and no shadow rays are allowed. What would be your importance-sampling strategy, and why at the first and second intersection points for a given path? [4 marks]
- (c) For a given scene, assume the rendering equation can be reduced to  $L_o(\mathbf{x}, \vec{\omega}_o) = c \int_{H^2} V(\mathbf{x}, \vec{\omega}_i) \cos \theta_i d\vec{\omega}_i$  at surface points  $\mathbf{x}$ , where  $V$  is the (binary) visibility function.
- (i) What are the properties of the scene that make this reduction possible? [3 marks]
  - (ii) Assuming a single object and ignoring inter-reflections for that object, what is the gradient of  $L_o(\mathbf{x}, \vec{\omega}_o)$  with respect to  $\mathbf{x}$ ? [2 marks]



## 9 Further Human–Computer Interaction

- (a) Briefly explain, including quantitative estimates, how the information content of English text compares to the information rate that could theoretically be communicated using a computer keyboard. [2 marks]
- (b) How could the relationship you have described be used in a design to enter text either more quickly, or with less effort from the user? [2 marks]
- (c) Consider an operating system command line interpreter that interprets text commands on a probabilistic basis. Explain how each of the elements of Bayes' theorem (Posterior, Likelihood, Prior and Evidence) would be applied in the implementation of this interpreter. [8 marks]
- (d) Give one example of a well-known bias or heuristic that might be applied by users of a command line interpreter when they need to make rapid decisions. Explain, using Bayes theorem and/or information theory, what could make that strategy effective in the above scenario. [4 marks]
- (e) Explain how the attention investment model of abstraction use might relate to the use of more or less abstract strategies in the scenario you have described. [4 marks]

## 10 Further Human–Computer Interaction

This question relates to the design of an online tool for use by a local radio station, that can be used to more efficiently generate scripts for their bulletins. The station manager has asked you to investigate whether presenters could ask ChatGPT to automatically generate the evening weather report, which typically includes a summary of today’s weather, and forecasts for tomorrow and the rest of the week.

- (a) Sketch a user interface that might be used as the front end to a generative AI large language model, as an alternative to the standard conversational style of interaction with ChatGPT. [3 marks]
- (b) Add labels to refer to two specific parts of your sketch, and explain how the graphical properties of those elements correspond to their function within your design. [4 marks]
- (c) What analytic method could be used to predict whether your alternative design would be more time-efficient than the standard style of interaction with ChatGPT? [1 mark]
- (d) Based on the analytic method that you suggest, provide calculations for an estimated (within an order of magnitude) comparison of (i) the station manager’s suggestion to use ChatGPT, and (ii) the use of the user interface you have sketched, in a typical usage. [4 marks]
- (e) Describe an empirical approach that could be used to evaluate whether the estimates that you have calculated are justified. [4 marks]
- (f) Describe up to four further considerations, beyond time-efficiency, that might be relevant to the station manager’s decision to deploy such a tool. Make specific reference to topics from the HCI literature, e.g. from the lectures, notes or readings, but not media reports of AI. [4 marks]

**END OF PAPER**